Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2015/03/jls.htm
2015 Vol. 5 (S3), pp. 1368-1375/Borujerdi and Fereydooni
Research Article

SIMULATION OF UNDERGROUND WATER LEVEL USING SUPPORT VECTOR MACHINES(SVM) AND ITS COMPARISON WITH ADAPTIVE NEURO FUZZY INFERENCE SYSTEM(ANFIS) AND NEURAL WAVELET MODELS(WNN) (CASE STUDY: SHIRAZ PLAIN)

***Arash Borujerdi[1] and Mehrdad Fereydooni[2]**
[1]Department of Hydraulic Structures, Larestan Branch, Islamic Azad University, Larestan, Iran
[2]Department of Civil Engineering, Larestan Branch, Islamic Azad University, Larestan, Iran
**Author for Correspondence*

ABSTRACT
Human's excessive usage of underground waters for different consumptions has led to day-to-day reduction in volume of these resources. In the present paper, underground water level and precipitation data of Shiraz Plain were used to model underground water balance. The present paper dealt with scientific fundamentals of the three methods applied in the research i.e. SVM, ANFIS, and neural wavelet, their different techniques, and implementation methods. Also, the algorithms were implemented using available software packages and the acquired diagrams and models were illustrated; here, MATLAB software toolboxes were selected. Ultimately, monthly water balance of Shiraz was predicted using data of two monthly time-steps of precipitation and water level data for an interval of 19.5 years since March 1993 until September 2013. In the best case, the regression coefficients of 0.993, 0.986, and 0.767 were achieved for neural wavelet, ANFIS, and support vector machine (SVM) models, respectively. In addition, the following values of mean square error (MSE) were obtained respectively for neural wavelet model, ANFIS, and support vector machine (SVM): 0.0003, 0.0019, and 0.0058, all of which suggest a high accuracy. Neural wavelet model was the most successful while SVM model exhibited the poorest performance among the analyzed models.

Keywords: *Modeling, Underground Water, ANFIS, Support Machine Vector, Neural Wavelet*

INTRODUCTION
Excessive usage of underground waters by humans for different consumptions has resulted in exceeding decline of the volume of the respective resources. Inconsiderate and uncontrolled extraction of underground waters, which usually exceeds the natural feed, will bring about many detrimental consequences such as reduction of underground water level, dried wells, decline in water volume and/or drying of aqueducts, springs, and streams, water quality impairment, increase in pumping costs, ground subsidence, and reduction of crops.

The underground water level and precipitation data of Shiraz Plain were used in the current research to model the underground water level. Taking into account large population of Shiraz Plain residents and the fact that Shiraz City is situated on this plain, it is highly significant to carry out the present study and analyze the underground water which supplies a major portion of urban, industrial, and agricultural water requirements. Providing a suitable forecast might prove helpful to organizations and industries.

With corroboration of ability of artificial neural networks, fuzzy logic, and other adaptive methods of artificial intelligence in simulation of non-stationary and non-linear times series during the recent years, extensive research works have been conducted in different fields using these powerful tools all around the world. The respective models do not require long-term statistics for prediction of time series.

These methods have widespread applications in hydrological models and simulation of their related parameters thanks to relatively low computational cost, lack of necessity to estimate the parameters affecting the governing equations (such as hydraulic conductivity coefficient in Richards Equation), and simple nature of random models. Different types of these methods include artificial intelligence and support vector methods. Artificial intelligence embraces a variety of fuzzy logic and neural network techniques and their combinations with other mathematical algorithms.

The main idea of neural network originates from theories put forth by William James in 1890. The well-known neurologist Mcculloch and the prominent statistician Pitts in 1943 proposed a primary model of neuron. In 1949, Donald Hebb, the Canadian psychologist depicted instruction of neural networks from psychological standpoint in a book entitled "The Organization of Behavior". Eventually, Frank Rosenblatt, neuro-biologist of Kernel University, introduced the perceptron algorithm in 1957.

Daniel *et al.,* (1991) benefited from artificial neural networks as the pioneer in the scope of hydrological discussions associated with prediction of monthly water consumption and flood. Afterwards, along with other hydrological processes, utilization of artificial intelligence models expanded in prediction of underground water balance due to their simple applicability and high accuracy.

Jazaee *et al.,* (2008) forecast fluctuations of underground urban water level using Geographical Information System (GIS) and Artificial Neural Network (ANN) model. He concluded that: first, accuracy of artificial neural network improves using GIS data; second, artificial neural network models have high ability in prediction of underground water level. Third, the three training algorithms used in his research namely, LM, BR, and GDX nearly have the same level of accuracy in predictions. However, performance of LM training is slightly better than the two other ones.

Since advent of novel theory of wavelet transform in the field of mathematics and engineering sciences, utilization of wavelets has rapidly increased as a strong tool in decomposition and preprocessing of time series and its capabilities in improvement of prediction results have been confirmed. Though not very old, numerous researches have been carried out using wavelet theory during the recent years. Torrence and Campo (1998) analyzed wavelets and their capabilities in assessing spatial variations of time series as a practical tool. Woong and Valdes (2003) took advantage of hybrid model of artificial neural network and wavelet transforms for predicting droughts in Kansas River basin in Mexico. Implying to potentials of wavelet transform in decomposition of time series, they inferred that the non-linear and adaptive model of artificial neural network and wavelet transforms improved accuracy of drought prediction compared to artificial neural network model.

Mehdikhani *et al.,* (2006) used an adaptive model based on artificial neural network and wavelet transforms for drought prediction. He reached to the conclusion that the adaptive WNN model considerably improved accuracy of drought prediction compared to ANN model.

Adamowsky and Chen (2011) utilized 8-year statistical data of Chateauguay drainage basin in Canada and predicted underground water level using adaptive and non-linear model of artificial neural network and wavelet transforms. They compared results of underground water level predicted by the two respective methods. The results indicated that the adaptive model is more accurately capable of predicting underground water level. Nonetheless, they recommended further studies to be performed in order for better and more effective management of underground water resources.

Jang *et al.,* (1993) founded Adaptive Neuro-Fuzzy Inference System (ANFIS) via integrating lingual strength of fuzzy systems and training ability of neural networks (Chen *et al.,* 2010). Following completion of this adaptive model, its domain of application expanded and is still expanding in modeling of different processes such as hydrological processes.

Bazartseren *et al.,* (2003) predicted water levels of two rivers in Germany using ANN and ANFIS models. Simulation results indicated both models have high ability to predict water level compared to linear models. However, error of ANN model was smaller than ANFIS in both rivers.

Kouhestani *et al.,* (2011) predicted underground water level of Normab Plain of Golestan Province of Iran using neuro-fuzzy inference system model. Model inputs included geographical location of piezometers, underground water levels, amount of water inlet to each polygon, precipitation amount, and amount of water extraction from each piezometer. Modeling results demonstrated that the respective model has a suitable accuracy with negligible error and correlation coefficient of 0.99.

Amutha and Porchelvan (2011) forecast seasonal underground water level in Malatar Sub-basin of India using ANFIS and radial-based function (RBF) models. Input data of models were underground water level and precipitation in previous seasons. Research results suggested that ANFIS model has smaller error despite acceptable prediction accuracy of both models.

Asefa *et al.*, (2008) proposed a proper method for predicting seasonal and hourly flow rate using SVM model. Using equivalent water amount of snow and flow volume in the previous periods in the respective research, volumetric flow rate was estimated for 6-month and 24-hour time scales and the results were satisfactory.

Through a study, Yu *et al.*, (2010) predicted digital level of flood in a river in Taiwan using SVM model. They took advantage of network search optimization method for calibrating the model parameters. Simulation results manifested that the model properly predicts river water level for the next 1-6 hours.

Shivam *et al.*, (2012) deployed SVM and ANN models for simulation of climate changes in India. The results are indicative of the fact that SVM can be a suitable alternative for ANN in prediction of precipitation variations.

Wen *et al.*, (2014) compared different artificial intelligence methods in prediction of monthly stream flow. The results show that SVM has higher capability compared to ANN, ARIMA, and ANFIS in simulation of monthly stream flow.

Region of Study

The region of study is Shiraz Plain with a length of 40 km and varying width of 15-30 km situated between 29° 30` and 29° 39` northern latitudes and 52° 25` and 52° 44` eastern longitudes in southwest Iran (Figure 1). The the region has a surface area of approximately 261 km^2 and maximum and minimum altitudes from sea level are 1599 and 1455 meters, respectively. Its annual average temperature is 18° and average yearly precipitation is 337 mm. This plain is constrained by Derak Mountain to the west, Bamou, Sabzpoushan and Chehel Maqam mountains and BabaKouhi Mountain (part of Zagros Mountain Range) to the north.

Keeping in mind the large population of residents in Shiraz Plain and the fact that Shiraz City is located on this plain, it is greatly significant to carry out the current study and assess the level of underground water which supplies a large portion of urban, industrial, and agricultural water requirement. Providing a proper prediction can solve problems of organizations and industries (Shaghayeghian and Mahshid, 2011).

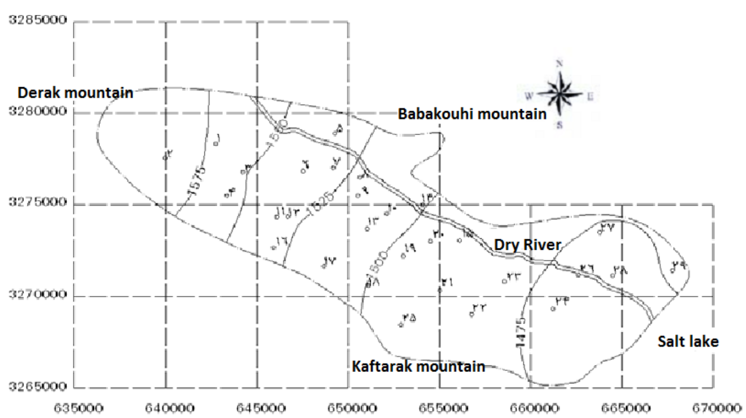

Figure 1: The region of study and locations of observation wells

1- *Neural Wavelet Technique*

Wavelets are a group of mathematical functions used for decomposing continuous signal into frequency components where resolution of each component equals its scale. Wavelet transform is decomposition of a function based on wavelet functions. Wavelets (known as daughter wavelets) are transferred and scaled samples of a function (mother wavelet) with finite length and an extremely attenuating frequency. There are two ways to combine wavelet and neural networks as implied in the literature:

The signal is first decomposed at several levels by means of discrete wavelet transform and its capability in signal analysis with different resolutions. Then, a feature vector is extracted for each perturbation using standard deviation of wavelet coefficients at each level, which is the basis for signal classification by neural network. In the respective method, wavelet is applied in the core of MLP neural network. As explained earlier, a neural network is composed of several neurons and each neuron has a transfer function at its output. In this method, transfer function of neuron undergoes variation. As such, first the time series is decomposed by wavelet toolbox and then its high frequencies are given to the neural network program as the main signal.

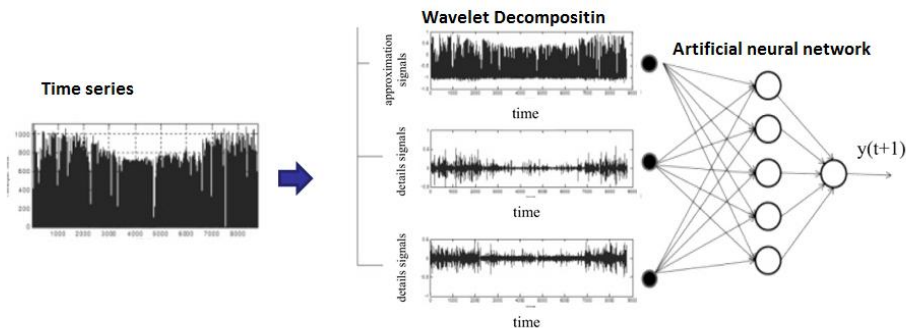

Figure 2: Neural wavelet network implimented in the present paper

Also, the responses of diagrams related to results of this simulation are illustrated in the following figures. It must be mentioned that the results of this simulation are worse than other methods in terms of R and MSE.

2- Adaptive Neuro-Fuzzy Inference System (Anfis) Technique

In this theory, the elements are evaluated in steps and continuously; the focus is on approximate reasoning. In approximate reasoning, it is attempted to acquire certain output from uncertain inputs. Results of this method depend on selection of fuzzy rules and fuzzy functions. In classic sets, there are only two sorts of membership between the element and set: membership and non-membership whereas fuzzy logic states include low, high, good, moderate, bad, and other uncertain states.

Adaptive neuro-fuzzy inference system (ANFIS) model was initially proposed by Jang (1993). The main framework of ANFIS is based on the concept of fuzzy inference system. This model has self-organization ability as the principal characteristic of artificial neural network. In fuzzy inference system, human knowledge or experience and inference process can be qualitatively described and analyzed via if-then rules but quantitative analysis cannot be performed. Artificial neural networks have the abilities of automatic learning, organizing, and adapting but are not qualitative and the inference process is not comparable. Jang merged advantages of fuzzy inference system and introduced ANFIS method. As a result, ANFIS simultaneously enjoys capabilities of both self-organizing structure and being qualitative. Advantages of ANFIS model can be summarized as follows: combining rules and data of real world performance, selection and optimization of rules and input/output variables during training stage, imitating human decision-making process, and rapid computation using fuzzy operation.

3- Support Vector Machine (SVM)

Support Vector Machines (SVMs) is one of the supervised learning methods utilized for classification and regression. The primary SVM algorithm was innovated in 1963 by Vladimir Vapnik and later generalized for non-linear state in 1995 by Vapnik and Corinna Cortes. This method is among the relatively new techniques which has exhibited good efficiency for classification during the recent years compared to older methods such as perceptron neural networks. The operation of SVM classifier is based on linear classification of data, and, a line with larger confidence margin is tried to be chosen in linear division of data. Quadratic Programming (QP) methods which are famous techniques in solving constrained problems are employed to solve the problem of finding optimal line for data.

4- Analysis of Results

In implementation of different models with different modeling techniques, previous values of a variable are required for simulation, modeling, and prediction. These data in fact constitute input data of our modeling technique.

Here, four layouts were used for all three modeling techniques. These four layouts include:
1- First layout: using (P_{-1},w_{-1})
2- Second layout: using (w_{-1})
3- Third layout: using (w_{-1},w_{-2})
4- Fourth layout: using (P_{-1},w_{-1},w_{-2},P_{-2})

In the present paper, three models i.e. neural wavelet, ANFIS, and SVM were run for simulation using the mentioned layouts and with the aid of MATLAB software. The data, provided by Regional Water Organization, were underground water level and precipitation amount in different months of the year since March 1993 until September 2013. These data are averages of the values measured in 29 wells throughout the region.

6-1- Results of Neural Wavelet Technique

Response of diagrams related to this simulation are included in the following figures. It is worth mentioning that the results of these simulations are superior to other methods in terms of R and MSE. Results of its analysis are included in Table 1. These results were acquired for the second layout with R=0.9946 and MSE=0.00033, as the best result of all techniques.

Table 1: Results related to 4 layouts of neural wavelet technique

Fourth	Third	Second	First	Layout
0.00082	0.0019	0.000339	0.0059	**MSE**
0.9901	0.9818	0.9946	0.9873	**R**

The results of this simulation are also shown in the figures below and it worths noting that this results for R and MSE are illustrated by MATLAB.

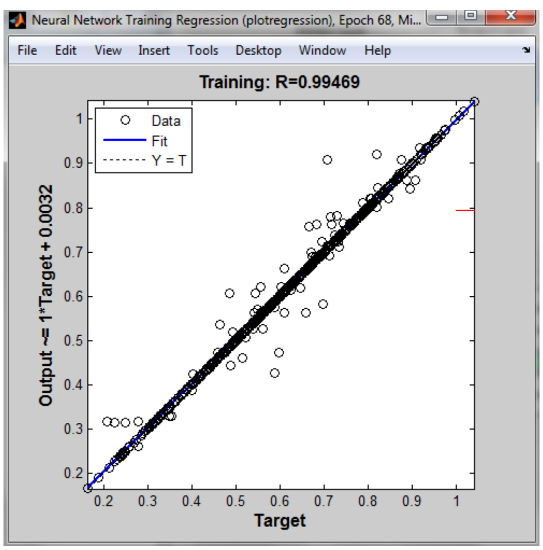

Figure 3: Correlation diagram in the second layout of neural wavelet

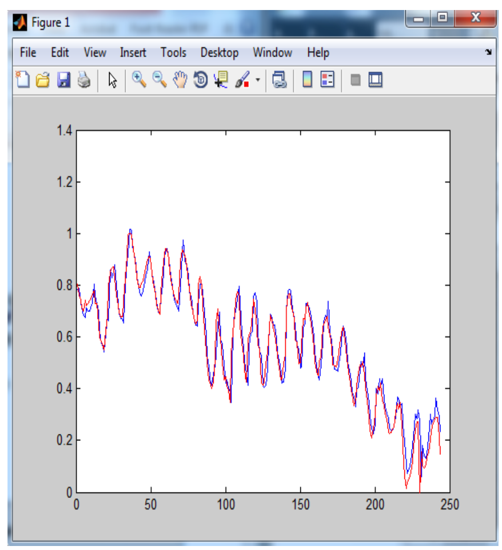

Figure 4: Test and objective functions of the first layout in neural network system

6-2- Anfis Technique Results

Simulation was also performed for adaptive neuro-fuzzy inference system (ANFIS) technique in the present research, which yielded very good and desirable results. As observed in the following table (Table 2), the best results of this technique are achieved in the second layout i.e. using (w_{-1}) where only one backward monthly time-step of the relevant dataset is used in simulation input; the data have been normalized in the previous stage. The best results related to this technique which were slightly poorer than neural wavelet technique belonged to the second layout with R=0.9818 and MSE=0.0032.

Table 2: The results related to 4 layouts of ANFIS technique

Fourth	Third	Second	First	Layout
0.0024	0.0019	0.0032	0.0024	**MSE**
0.9767	0.9688	0.9818	0.9781	**R**

The respective best results are illustrated in Figures 5 and 6, respectively related to regression and mean square error in the second simulation layout using ANFIS.

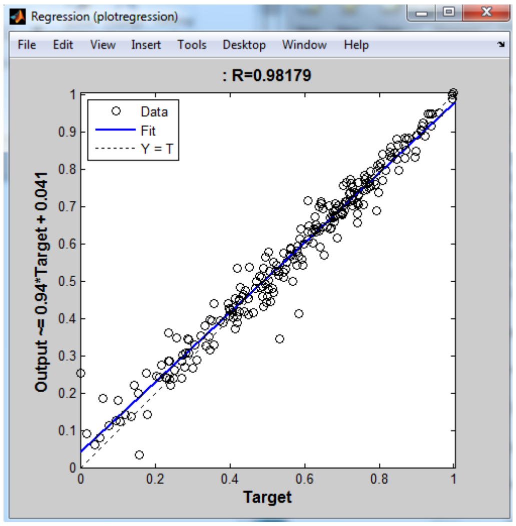

Figure 5: Correlation results related to second ANFIS layout

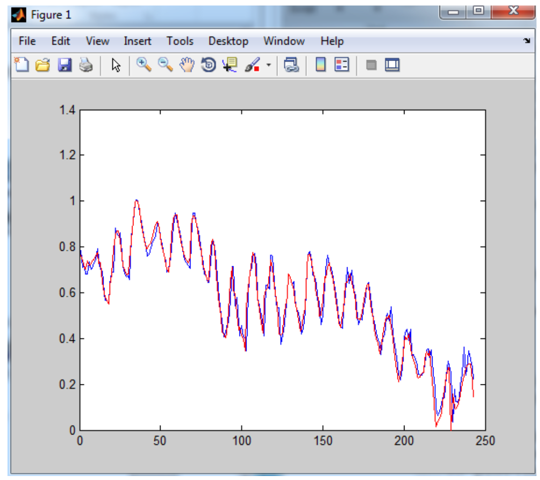

Figure 6: Diagram of test and objective functions for the third layout in ANFIS system

6-1- Results of Support Vector Machine Technique

In our assumptions, support vector machine model technique was probably supposed to yield superior results for prediction due to its novelty but the assumption turned out to be wrong and the results of SVM technique were poorer than three other techniques.

This issue can be observed in Table (3) where 4 simulation layouts did not yield a result better than 0.7673. The best executed layout is the first one i.e. using (P_{-1}, w_{-1}).

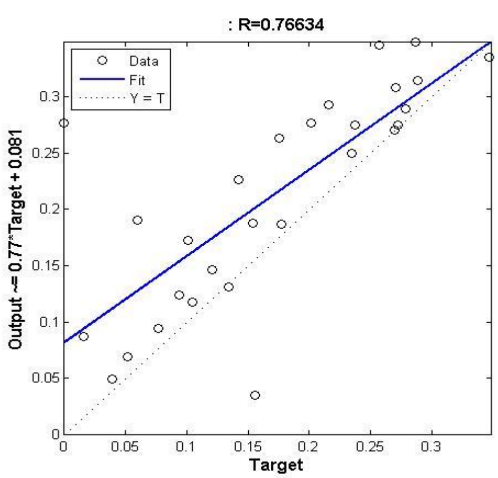

Figure 7: Regression results of the first layout in SVM modeling technique

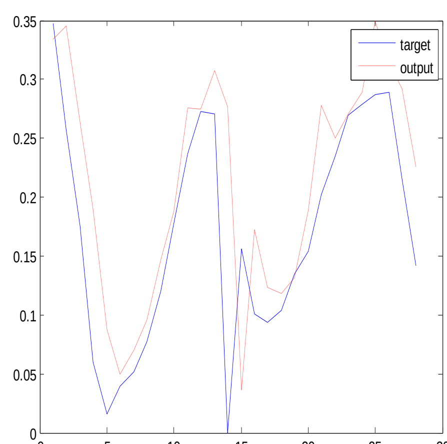

Figure 8: Test and objective functions of the first layout in SVM modeling technique

Table 3: Results related to 4 layouts of support vector machine technique

Fourth	Third	Second	First	Layout
0.0068	0.0082	0.0059	0.0058	**MSE**
0.7169	0.6671	0.6917	0.7673	**R**

These results, best of which had a mean square error of 0.0058, are illustrated in the above figures (respectively, regression and MSE diagrams) (figure 7, 8).
The authors managed to propose and compare different artificial intelligence models. Each of models had been assessed separately in former researches in hydrological applications. The results can be observed in Figures 2 through 7.
Results of all modeling techniques were nearly acceptable, and in fact, iteration level of modeling technique occasionally reflects accuracy of modeling technique. However, it was assumed in the present research that the results belonging to SVM shall be better than others. This assumption was not confirmed during simulation and the most desirable results belonged to neural wavelet technique.

Table 4: Comparison of the best result for three modeling techniques

SVM	ANFIS	Neural Wavelet	Modeling Technique
07673	0.9818	0.9933	**Best R**
0.0059	0.0019	0.0003396	**Best MSE**
First	Third	Second	**Best Layout**

As seen, results of three modeling techniques show that each of layouts generates superior response in one of the modeling techniques. It cannot be therefore certainly asserted which layout is the best and all three models need to be tested to achieve the best results. The following table demonstrates in which modeling technique superior results have been acquired from each layout.

Table 5: Comparison of the best results for each of 4 simulation layouts

Fourth	Third	Second	First	Layout
0.00082	0.0019	0.000339	0.0059	**MSE**
0.9901	0.9818	0.9933	0.7673	**R**
Neural Wavelet	ANFIS	Neural Wavelet	SVM	**Modeling Technique**

Conclusion

In the present paper, the underground water level and precipitation data of Shiraz Plain were used to model underground water balance. This study and assessment of underground water balance is greatly significant taking into account large population residing in this plain including Shiraz City's inhabitants and also the fact that these water resources supply a large portion of water requirement in urban, industrial, and agricultural sectors. Providing an appropriate prediction might be helpful to organizations and industries. The authors managed to evaluate simultaneously different artificial intelligence models which, in general, had been separately assessed in former research works in hydrological applications. Here, in addition to modeling, the different methods were compared and the superior method was introduced. Using three methods of neural wavelet, adaptive neuro-fuzzy inference system (ANFIS), and support machine vector (SVM), underground water level for the next month was modeled and predicted. The simulations were run with four layouts each of which yielded the best result in different models. Overall, neural wavelet model with R=0.993 was superior to other methods.

REFRENCES

Aqil M, Kita I and Yano A *et al.*, (2007). A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *Journal of Hydrology* **337** 22–34.

Chang CC and Lin CJ (2011). LIBSVM: a Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology* **2**(3) 1–27, Available: http:// www.csie.ntu.edu.tw/ cjlin/libsvm.

Chau KW, Wu CL and Li YS (2005). Comparison of several flood forecasting models in Yangtze River. *Journal of Hydrologic Engineering* **10** 485–491.

Chen ST and Yu PS (2007). Pruning of support vector networks on flood forecasting. *Journal of Hydrology* **347** 67–78.

Chen ST, Yu PS and Tang YH (2010). Statistical downscaling of daily precipitation using support vector machines and multivariate analysis. *Journal of Hydrology* **385** 13–22.

Daniel AR and Chen AA (1991). Stochastic simulation and forecasting of hourly average wind speed sequences in Jamaica. *Solar Energy* **46**(1) 1-11.

Dash NB, Panada SN, Remeson R and Sahoo N (2010). Hybrid neural modeling for groundwater level prediction. *Neural Computing and Applications*, DOI 10.1007/s00521-010-0360-1.

Gottardi G and Venutelli M (2007). Mathematical development and verification of a non-orthogonal finitevolume model for groundwater flowapplications. *Advances in Water Resources* **30**(1) 29-42.

Jazaee Farhad (2008). Prediction of Urban Underground Water Level by means of Geographic Information System (GIS) and Artificial Neural Network, Master Thesis at Sharif University.

Kholghi M and Hosseini SM (2008). Comparison of Groundwater Level Estimation Using Neuro-fuzzy and Ordinary Kriging. *Environmental Monitoring and Assessment*, DOI 10.1007/s10666-008-9174-2.

Kim Tae-Woong and Juan B Valdés (2003). Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *Journal of Hydrologic Engineering* **8.6** 319-328.

Rees I, Masters I, Malan AG and Lewis RW (2004). An edge-base finite volume scheme for saturatedunsaturated groundwater flow. *Computer Methods in Applied Mechanics and Engineering* **19**(42-44) 4741-4759.

Reghunath R, Murthy TRS Urthy and Raghavan BR (2005). Time series analysis to monitor and assess water resources: a moving average approach. *Environmental Monitoring and Assessment* **109** 65–72.

Shaghayeghian Mahmoudreza and Mahshid Shaghaghian (2011). Comparison of Time Modeling (Prediction) of Underground Water Level using Fuzzy Logic and Basic Methods of Time Series. Analysis (Case Study: Shiraz Plain), *6th National Congress of Civil Engineering, Semnan, University of Semnan.*

Tuckfield RC (1994). Estimating an appropriate sampling frequency for monitoring ground water well contamination", paper presented at International Nuclear Materials Management (INMM) Annual Meeting,Naples.